

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Briskin, Michael J.
Ringler, Douglas J.
Picarella, Dominic
Newman, Walter
- (ii) TITLE OF INVENTION: Mucosal Vascular Addressins and Uses
Thereof
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 - (B) STREET: Two Militia Drive
 - (C) CITY: Lexington
 - (D) STATE: Massachusetts
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 02421-4799
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/875,849
 - (B) FILING DATE: 08-SEP-1997
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: PCT/US96/02153
 - (B) FILING DATE: 12-FEB-1996
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/523,004
 - (B) FILING DATE: 01-SEP-1995
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/386,857
 - (B) FILING DATE: 10-FEB-1995
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Brook, David E.
 - (B) REGISTRATION NUMBER: 22,592
 - (C) REFERENCE/DOCKET NUMBER: LKS94-04A2
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 781-861-6240
 - (B) TELEFAX: 781-861-9540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1624 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1218

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAT TTC GGA CTG GCC CTC CTG CTG GCG GGG CTT CTG GGG CTC CTC	48
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu	
1 5 10 15	
CTC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG	96
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu	
20 25 30	
CCG GTG GTG GCC GTG GCC TTG GGC GCC TCG CGC CAG CTC ACC TGC CGC	144
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg	
35 40 45	
CTG GCC TGC GCG GAC CGC GGG GCC TCG GTG CAG TGG CGG GGC CTG GAC	192
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp	
50 55 60	
ACC AGC CTG GGC GCG GTG CAG TCG GAC ACG GGC CGC AGC GTC CTC ACC	240
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr	
65 70 75 80	
GTG CGC AAC GCC TCG CTG TCG GCG GCC GGG ACC CGC GTG TGC GTG GGC	288
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly	
85 90 95	
TCC TGC GGG GGC CGC ACC TTC CAG CAC ACC GTG CAG CTC CTT GTG TAC	336
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr	
100 105 110	
GCC TTC CCG GAC CAG CTG ACC GTC TCC CCA GCA GCC CTG GTG CCT GGT	384
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly	
115 120 125	
GAC CCG GAG GTG GCC TGT ACG GCC CAC AAA GTC ACG CCC GTG GAC CCC	432
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro	
130 135 140	
AAC GCG CTC TCC TTC TCC CTG CTC GTC GGG GGC CAG GAA CTG GAG GGG	480
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly	
145 150 155 160	

GCG CAA GCC CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG CCC CAG	528
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln	
165 170 175	
GGG GAC GAG GAC GTG CTG TTC AGG GTG ACA GAG CGC TGG CGG CTG CCG	576
Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro	
180 185 190	
CCC CTG GGG ACC CCT GTC CCG CCC GCC CTC TAC TGC CAG GCC ACG ATG	624
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met	
195 200 205	
AGG CTG CCT GGC TTG GAG CTC AGC CAC CGC CAG GCC ATC CCC GTC CTG	672
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu	
210 215 220	
CAC AGC CCG ACC TCC CCG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	720
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
225 230 235 240	
CCC AAC ACC ACC TCC CCG GAG TCT CCC GAC ACC ACC TCC CCG GAG TCT	768
Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser	
245 250 255	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CAG GAG CCT	816
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro	
260 265 270	
CCC GAC ACC ACC TCC CAG GAG CCT CCC GAC ACC ACC TCC CCG GAG CCT	864
Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro	
275 280 285	
CCC GAC AAG ACC TCC CCG GAG CCC GCC CCC CAG CAG GGC TCC ACA CAC	912
Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His	
290 295 300	
ACC CCC AGG AGC CCA GGC TCC ACC AGG ACT CGC CGC CCT GAG ATC TCC	960
Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser	
305 310 315 320	
CAG GCT GGG CCC ACG CAG GGA GAA GTG ATC CCA ACA GGC TCG TCC AAA	1008
Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys	
325 330 335	
CCT GCG GGT GAC CAG CTG CCC GCG GCT CTG TGG ACC AGC AGT GCG GTG	1056
Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val	
340 345 350	
CTG GGA CTG CTG CTC CTG GCC TTG CCC ACG TAT CAC CTC TGG AAA CGC	1104
Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg	
355 360 365	
TGC CGG CAC CTG GCT GAG GAC GAC ACC CAC CCA CCA GCT TCT CTG AGG	1152
Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg	
370 375 380	
CTT CTG CCC CAG GTG TCG GCC TGG GCT GGG TTA AGG GGG ACC GGC CAG	1200
Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln	
385 390 395 400	

21
cut

GTC GGG ATC AGC CCC TCC TGAGTGGCCA GCCTTTCCCC CTGTGAAAGC	1248
Val Gly Ile Ser Pro Ser	
405	
AAAATAGCTT GGACCCCTTC AAGTTGAGAA CTGGTCAGGG CAAACCTGCC TCCCATTCTA	1308
CTCAAAGTCA TCCCTCTGCT CACAGAGATG GATGCATGTT CTGATTGCCT CTTTGGAGAA	1368
GCTCATCAGA AACTCAAAAG AAGGCCACTG TTTGTCTCAC CTACCCATGA CCTGAAGCCC	1428
CTCCCTGAGT GGTCCCCACC TTTCTGGACG GAACCACGTA CTTTTTACAT ACATTGATTC	1488
ATGTCTCACG TCTCCCTAAA AATGCGTAAG ACCAAGCTGT GCCCTGACCA CCCTGGGCCC	1548
CTGTCGTCAG GACCTCCTGA GGCTTTGGCA AATAAACCTC CTAAAATGAT AAAAAAAAAA	1608
AAAAAAAAAA AAAAAA	1624

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 406 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
1 5 10 15

Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu
20 25 30

Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg
35 40 45

Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp
50 55 60

Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr
65 70 75 80

Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly
85 90 95

Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr
100 105 110

Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly
115 120 125

Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro
130 135 140

Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly
 145 150 155 160

Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln
 165 170 175

Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro
 180 185 190

Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met
 195 200 205

Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu
 210 215 220

His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro
 225 230 235 240

Pro Asn Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Pro Glu Ser
 245 250 255

Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Gln Glu Pro
 260 265 270

Pro Asp Thr Thr Ser Gln Glu Pro Pro Asp Thr Thr Ser Pro Glu Pro
 275 280 285

Pro Asp Lys Thr Ser Pro Glu Pro Ala Pro Gln Gln Gly Ser Thr His
 290 295 300

Thr Pro Arg Ser Pro Gly Ser Thr Arg Thr Arg Arg Pro Glu Ile Ser
 305 310 315 320

Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly Ser Ser Lys
 325 330 335

Pro Ala Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser Ser Ala Val
 340 345 350

Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu Trp Lys Arg
 355 360 365

Cys Arg His Leu Ala Glu Asp Asp Thr His Pro Pro Ala Ser Leu Arg
 370 375 380

Leu Leu Pro Gln Val Ser Ala Trp Ala Gly Leu Arg Gly Thr Gly Gln
 385 390 395 400

Val Gly Ile Ser Pro Ser
 405

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1539 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1146

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GAT TTC GGA CTG GCC CTC CTG CTG GCG GGG CTT CTG GGG CTC CTC	48
Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu	
1 5 10 15	
CTC GGC CAG TCC CTC CAG GTG AAG CCC CTG CAG GTG GAG CCC CCG GAG	96
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu	
20 25 30	
CCG GTG GTG GCC GTG GCC TTG GGC GCC TCG CGC CAG CTC ACC TGC CGC	144
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg	
35 40 45	
CTG GCC TGC GCG GAC CGC GGG GCC TCG GTG CAG TGG CGG GGC CTG GAC	192
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp	
50 55 60	
ACC AGC CTG GGC GCG GTG CAG TCG GAC ACG GGC CGC AGC GTC CTC ACC	240
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr	
65 70 75 80	
GTG CGC AAC GCC TCG CTG TCG GCG GCC GGG ACC CGC GTG TGC GTG GGC	288
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly	
85 90 95	
TCC TGC GGG GGC CGC ACC TTC CAG CAC ACC GTG CAG CTC CTT GTG TAC	336
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr	
100 105 110	
GCC TTC CCG GAC CAG CTG ACC GTC TCC CCA GCA GCC CTG GTG CCT GGT	384
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly	
115 120 125	
GAC CCG GAG GTG GCC TGT ACG GCC CAC AAA GTC ACG CCC GTG GAC CCC	432
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro	
130 135 140	
AAC GCG CTC TCC TTC TCC CTG CTC GTC GGG GGC CAG GAA CTG GAG GGG	480
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly	
145 150 155 160	
GCG CAA GCC CTG GGC CCG GAG GTG CAG GAG GAG GAG GAG GAG CCC CAG	528
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln	
165 170 175	
GGG GAC GAG GAC GTG CTG TTC AGG GTG ACA GAG CGC TGG CGG CTG CCG	576
Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro	
180 185 190	

CCC CTG GGG ACC CCT GTC CCG CCC GCC CTC TAC TGC CAG GCC ACG ATG	624
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met	
195 200 205	
AGG CTG CCT GGC TTG GAG CTC AGC CAC CGC CAG GCC ATC CCC GTC CTG	672
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu	
210 215 220	
CAC AGC CCG ACC TCC CCG GAG CCT CCC GAC ACC ACC TCC CCG GAG TCT	720
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Ser	
225 230 235 240	
CCC GAC ACC ACC TCC CCG GAG TCT CCC GAC ACC ACC TCC CAG GAG CCT	768
Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro	
245 250 255	
CCC GAC ACC ACC TCC CCG GAG CCT CCC GAC AAG ACC TCC CCG GAG CCC	816
Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro	
260 265 270	
GCC CCC CAG CAG GGC TCC ACA CAC ACC CCC AGG AGC CCA GGC TCC ACC	864
Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr	
275 280 285	
AGG ACT CGC CGC CCT GAG ATC TCC CAG GCT GGG CCC ACG CAG GGA GAA	912
Arg Thr Arg Arg Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu	
290 295 300	
GTG ATC CCA ACA GGC TCG TCC AAA CCT GCG GGT GAC CAG CTG CCC GCG	960
Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala	
305 310 315 320	
GCT CTG TGG ACC AGC AGT GCG GTG CTG GGA CTG CTG CTC CTG GCC TTG	1008
Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Ala Leu	
325 330 335	
CCC ACC TAT CAC CTC TGG AAA CGC TGC CGG CAC CTG GCT GAG GAC GAC	1056
Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp	
340 345 350	
ACC CAC CCA CCA GCT TCT CTG AGG CTT CTG CCC CAG GTG TCG GCC TGG	1104
Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp	
355 360 365	
GCT GGG TTA AGG GGG ACC GGC CAG GTC GGG ATC AGC CCC TCC	1146
Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser	
370 375 380	
TGAGTGGCCA GCCTTTCCCC CTGTGAAAGC AAAATAGCTT GGACCCCTTC AAGTTGAGAA	1206
CTGGTCAGGG CAAACCTGCC TCCCATTCTA CTCAAAGTCA TCCCTCTGTT CACAGAGATG	1266
GATGCATGTT CTGATTGCCT CTTTGGAGAA GCTCATCAGA AACTCAAAAG AAGGCCACTG	1326
TTTGTCTCAC CTACCCATGA CCTGAAGCCC CTCCCTGAGT GGTCCCCACC TTTCTGGACG	1386
GAACCACGTA CTTTTTACAT ACATTGATTC ATGTCTCACG TCTCCCTAAA AATGCGTAAG	1446
ACCAAGCTGT GCCCTGACCA CCCTGGGCCC CTGTCGTCAG GACCTCCTGA GGCTTTGGCA	1506

AATAAACCTC CTAAATGAA AAAAAAAAAA AAA

1539

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 382 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asp Phe Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
1 5 10 15
Leu Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu Pro Pro Glu
20 25 30
Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu Thr Cys Arg
35 40 45
Leu Ala Cys Ala Asp Arg Gly Ala Ser Val Gln Trp Arg Gly Leu Asp
50 55 60
Thr Ser Leu Gly Ala Val Gln Ser Asp Thr Gly Arg Ser Val Leu Thr
65 70 75 80
Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val Cys Val Gly
85 90 95
Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Gln Leu Leu Val Tyr
100 105 110
Ala Phe Pro Asp Gln Leu Thr Val Ser Pro Ala Ala Leu Val Pro Gly
115 120 125
Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro Val Asp Pro
130 135 140
Asn Ala Leu Ser Phe Ser Leu Leu Val Gly Gly Gln Glu Leu Glu Gly
145 150 155 160
Ala Gln Ala Leu Gly Pro Glu Val Gln Glu Glu Glu Glu Glu Pro Gln
165 170 175
Gly Asp Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp Arg Leu Pro
180 185 190
Pro Leu Gly Thr Pro Val Pro Pro Ala Leu Tyr Cys Gln Ala Thr Met
195 200 205
Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile Pro Val Leu
210 215 220
His Ser Pro Thr Ser Pro Glu Pro Pro Asp Thr Thr Ser Pro Glu Ser
225 230 235 240

Pro Asp Thr Thr Ser Pro Glu Ser Pro Asp Thr Thr Ser Gln Glu Pro
245 250 255

Pro Asp Thr Thr Ser Pro Glu Pro Pro Asp Lys Thr Ser Pro Glu Pro
260 265 270

Ala Pro Gln Gln Gly Ser Thr His Thr Pro Arg Ser Pro Gly Ser Thr
275 280 285

Arg Thr Arg Arg Pro Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu
290 295 300

Val Ile Pro Thr Gly Ser Ser Lys Pro Ala Gly Asp Gln Leu Pro Ala
305 310 315 320

Ala Leu Trp Thr Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu
325 330 335

Pro Thr Tyr His Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Asp
340 345 350

Thr His Pro Pro Ala Ser Leu Arg Leu Leu Pro Gln Val Ser Ala Trp
355 360 365

Ala Gly Leu Arg Gly Thr Gly Gln Val Gly Ile Ser Pro Ser
370 375 380

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1721 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 4..1038

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGC	ATG	GAT	CGG	GGC	CTG	GCC	CTC	CTG	CTG	GCG	GGG	CTT	CTG	GGG	CTC	48
Met	Asp	Arg	Gly	Leu	Ala	Leu	Leu	Leu	Ala	Gly	Leu	Leu	Gly	Leu		
1				5					10					15		
CTC	CAG	CCG	GGC	TGC	GGC	CAG	TCC	CTC	CAG	GTG	AAG	CCC	CTG	CAG	GTG	96
Leu	Gln	Pro	Gly	Cys	Gly	Gln	Ser	Leu	Gln	Val	Lys	Pro	Leu	Gln	Val	
				20					25					30		
GAG	CCC	CCG	GAG	CCG	GTG	GTG	GCC	GTG	GCC	CTG	GGC	GCC	TCT	CGC	CAG	144
Glu	Pro	Pro	Glu	Pro	Val	Val	Ala	Val	Ala	Leu	Gly	Ala	Ser	Arg	Gln	
			35					40					45			

CTC	ACC	TGC	CGC	CTG	GAC	TGC	GCG	GAC	CGC	GGG	GCC	ACG	GTG	CAG	TGG	192
Leu	Thr	Cys	Arg	Leu	Asp	Cys	Ala	Asp	Arg	Gly	Ala	Thr	Val	Gln	Trp	
		50					55					60				
CGG	GGC	CTG	GAC	ACC	AGC	CTG	GGC	GCG	GTG	CAG	TCG	GAC	GCG	GGC	CGC	240
Arg	Gly	Leu	Asp	Thr	Ser	Leu	Gly	Ala	Val	Gln	Ser	Asp	Ala	Gly	Arg	
	65					70					75					
AGC	GTC	CTC	ACC	GTG	CGC	AAC	GCC	TCG	CTG	TCG	GCG	GCC	GGG	ACC	CGT	288
Ser	Val	Leu	Thr	Val	Arg	Asn	Ala	Ser	Leu	Ser	Ala	Ala	Gly	Thr	Arg	
	80				85				90						95	
GTG	TGC	GTG	GGC	TCC	TGC	GGG	GGC	CGC	ACC	TTC	CAG	CAC	ACC	GTG	CGG	336
Val	Cys	Val	Gly	Ser	Cys	Gly	Gly	Arg	Thr	Phe	Gln	His	Thr	Val	Arg	
				100					105					110		
CTC	CTT	GTG	TAC	GCC	TTC	CCG	GAC	CAG	CTG	ACC	ATC	TCC	CCG	GCA	GCC	384
Leu	Leu	Val	Tyr	Ala	Phe	Pro	Asp	Gln	Leu	Thr	Ile	Ser	Pro	Ala	Ala	
			115					120					125			
CTG	GTG	CCT	GGT	GAC	CCG	GAG	GTG	GCC	TGT	ACG	GCC	CAC	AAA	GTC	ACG	432
Leu	Val	Pro	Gly	Asp	Pro	Glu	Val	Ala	Cys	Thr	Ala	His	Lys	Val	Thr	
		130					135					140				
CCT	GTG	GAC	CCC	AAT	GCG	CTC	TCC	TTC	TCC	CTG	CTC	CTG	GGG	GAC	CAG	480
Pro	Val	Asp	Pro	Asn	Ala	Leu	Ser	Phe	Ser	Leu	Leu	Leu	Gly	Asp	Gln	
		145				150					155					
GAA	CTG	GAG	GGG	GCC	CAG	GCT	CTG	GGC	CCG	GAG	GTG	GAG	GAG	GAG	GAG	528
Glu	Leu	Glu	Gly	Ala	Gln	Ala	Leu	Gly	Pro	Glu	Val	Glu	Glu	Glu	Glu	
	160				165					170					175	
GAG	GAG	CCC	CAG	GAG	GAG	GAG	GAC	GTG	CTG	TTC	AGG	GTG	ACA	GAG	CGC	576
Glu	Glu	Pro	Gln	Glu	Glu	Glu	Asp	Val	Leu	Phe	Arg	Val	Thr	Glu	Arg	
				180					185					190		
TGG	CGG	CTG	CCG	ACC	CTG	GCA	ACC	CCT	GTC	CTG	CCC	GCG	CTC	TAC	TGC	624
Trp	Arg	Leu	Pro	Thr	Leu	Ala	Thr	Pro	Val	Leu	Pro	Ala	Leu	Tyr	Cys	
			195					200					205			
CAG	GCC	ACG	ATG	AGG	CTG	CCT	GGC	TTG	GAG	CTC	AGC	CAC	CGC	CAG	GCC	672
Gln	Ala	Thr	Met	Arg	Leu	Pro	Gly	Leu	Glu	Leu	Ser	His	Arg	Gln	Ala	
		210					215					220				
ATC	CCG	GTC	CTG	CAC	GGC	CCG	ACC	TCC	CGG	GAG	CCC	CCC	GAC	ACG	ACC	720
Ile	Pro	Val	Leu	His	Gly	Pro	Thr	Ser	Arg	Glu	Pro	Pro	Asp	Thr	Thr	
		225				230					235					
TCC	CCG	GAA	CCC	CGG	GCC	GCG	ACC	TCC	CCG	GAG	ACC	ACC	CCC	CAG	CAG	768
Ser	Pro	Glu	Pro	Arg	Ala	Ala	Thr	Ser	Pro	Glu	Thr	Thr	Pro	Gln	Gln	
	240				245					250				255		
GGC	TCC	ACA	CAC	AGC	CCC	AGG	AGC	CCG	GGC	TCT	ACC	AGG	ACT	TGC	CGC	816
Gly	Ser	Thr	His	Ser	Pro	Arg	Ser	Pro	Gly	Ser	Thr	Arg	Thr	Cys	Arg	
				260					265					270		
CCT	GAG	ATC	TCC	CAG	GCT	GGG	CCC	ACG	CAG	GGA	GAA	GTG	ATC	CCA	ACA	864
Pro	Glu	Ile	Ser	Gln	Ala	Gly	Pro	Thr	Gln	Gly	Glu	Val	Ile	Pro	Thr	
			275					280					285			

GGC TCG TCC AAA CCT ACG GGT GAC CAG CTG CCC GCG GCT CTG TGG ACC 912
 Gly Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr
 290 295 300

AGC AGT GCG GTG CTG GGA CTG CTG CTC CTG GCT TTG CCC ACC TAC CAC 960
 Ser Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His
 305 310 315

CTC TGG AAA CGT TGC CGG CAC CTG GCT GAG GAC GGC GCC CAC CCA CCA 1008
 Leu Trp Lys Arg Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro
 320 325 330 335

GCT TCT CTG AGT AGC CAG CCC TTC CCC CTG TGAAGGGAAA ATAGGTTGGA 1058
 Ala Ser Leu Ser Ser Gln Pro Phe Pro Leu
 340 345

CCCCTTCAAG CTGAGAACTG GTCGGGGCAA ACCTGCCTCC CATTCTATTC AAAGTCATCG 1118

CTCTGGTCAC AGAGAGGGAC GCACATTCTG ATTGCCTCCT TTGGAAAGGC TCATCAGAAA 1178

CTCAAAAGAA GGTGATCGTT TGTCCCGCCT ACCCGTGACC TGGAAGCCCC CGCCCCGCTC 1238

GAGTGACCCC TGACTTTCTG GACGGAACCA ACGTACTTCT TACATATATT GATTCATGTG 1298

TCATATCTCC CTAAAATGCG TAAAACCAGC TGTGCCCCGA CCACCTTGGG CCCCTGCCAT 1358

CAGGACCTCC TGAGGCTTTG GCAAATAAAC CTCCTAAAAG GATAGAAACT GAAACTTGTG 1418

GCCGGGCGCG GTGGCTCAAG CCTGTAATCC CAGCACTTTG GGAGGCCGAG GTGGGTGGAT 1478

CACGAGGTCA GGAGATCGAG ACCATCCTGG CTAACCCGTG AAACCCCGTC TCTACTAAAA 1538

AAATACAAAA ATTAGCCGGG AGCGGTGGCG GGCGCCTGTA GTCCCAGCTA CTCGGGAGGC 1598

TGAGGCAGGA GAATGGCGTG AACCCGGGAG GCGGAGCTTG CAGTGAGCTG AGATCCGGCC 1658

ACTGCACTCC AGCCTGGGGG ACAGAGCGAG ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA 1718

AAA 1721

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asp Arg Gly Leu Ala Leu Leu Leu Ala Gly Leu Leu Gly Leu Leu
 1 5 10 15

Gln Pro Gly Cys Gly Gln Ser Leu Gln Val Lys Pro Leu Gln Val Glu
 20 25 30

Pro Pro Glu Pro Val Val Ala Val Ala Leu Gly Ala Ser Arg Gln Leu
35 40 45

Thr Cys Arg Leu Asp Cys Ala Asp Arg Gly Ala Thr Val Gln Trp Arg
50 55 60

Gly Leu Asp Thr Ser Leu Gly Ala Val Gln Ser Asp Ala Gly Arg Ser
65 70 75 80

Val Leu Thr Val Arg Asn Ala Ser Leu Ser Ala Ala Gly Thr Arg Val
85 90 95

Cys Val Gly Ser Cys Gly Gly Arg Thr Phe Gln His Thr Val Arg Leu
100 105 110

Leu Val Tyr Ala Phe Pro Asp Gln Leu Thr Ile Ser Pro Ala Ala Leu
115 120 125

Val Pro Gly Asp Pro Glu Val Ala Cys Thr Ala His Lys Val Thr Pro
130 135 140

Val Asp Pro Asn Ala Leu Ser Phe Ser Leu Leu Leu Gly Asp Gln Glu
145 150 155 160

Leu Glu Gly Ala Gln Ala Leu Gly Pro Glu Val Glu Glu Glu Glu Glu
165 170 175

Glu Pro Gln Glu Glu Glu Asp Val Leu Phe Arg Val Thr Glu Arg Trp
180 185 190

Arg Leu Pro Thr Leu Ala Thr Pro Val Leu Pro Ala Leu Tyr Cys Gln
195 200 205

Ala Thr Met Arg Leu Pro Gly Leu Glu Leu Ser His Arg Gln Ala Ile
210 215 220

Pro Val Leu His Gly Pro Thr Ser Arg Glu Pro Pro Asp Thr Thr Ser
225 230 235 240

Pro Glu Pro Arg Ala Ala Thr Ser Pro Glu Thr Thr Pro Gln Gln Gly
245 250 255

Ser Thr His Ser Pro Arg Ser Pro Gly Ser Thr Arg Thr Cys Arg Pro
260 265 270

Glu Ile Ser Gln Ala Gly Pro Thr Gln Gly Glu Val Ile Pro Thr Gly
275 280 285

Ser Ser Lys Pro Thr Gly Asp Gln Leu Pro Ala Ala Leu Trp Thr Ser
290 295 300

Ser Ala Val Leu Gly Leu Leu Leu Leu Ala Leu Pro Thr Tyr His Leu
305 310 315 320

Trp Lys Arg Cys Arg His Leu Ala Glu Asp Gly Ala His Pro Pro Ala
325 330 335

Ser Leu Ser Ser Gln Pro Phe Pro Leu
340 345

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 18 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCTACTGCC AGGCCACG

18

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCCTGGGAG ATCTCAGGG

19

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCACGATGA GGCTGCCTGG

20

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GTGGAGCCTG GGCTCCTGGG

20

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAGCTTCC ACCATGGATT TCGGACTGGC CC

32

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCGACTAGTG TCGGGCTGTG CAGGAC

26

(2) INFORMATION FOR SEQ ID NO:13:

- 21
4.4
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GGACTAGTGG TTTGGACGAG CCTGTTG

27

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Gly Leu Asp Thr Ser Leu
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Region
(B) LOCATION: 2
(D) OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Ile or Leu"

(ix) FEATURE:

(A) NAME/KEY: Region
(B) LOCATION: 3
(D) OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Asp or Glu"

(ix) FEATURE:

(A) NAME/KEY: Region
(B) LOCATION: 4
(D) OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Thr or Ser"

(ix) FEATURE:

(A) NAME/KEY: Region
(B) LOCATION: 5
(D) OTHER INFORMATION: /product= "OTHER"
/label= variable
/note= "Xaa = Pro or Ser"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gly Xaa Xaa Xaa Xaa Leu
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Gln Ile Asp Ser Pro Leu
1 5

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 5 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Leu Asp Thr Ser Leu
1 5